Development and Use of a Guideline Entry Wizard to Convert Text Clinical Practice Guidelines to a Relational Format

Joseph W. Hales, PhD, Cynthia S. Gadd, PhD, MBA, David F Lobach, MD, PhD, MS
Division of Medical Informatics, Duke University Medical Center, Durham, NC

Computerization of clinical practice guidelines (CPGs) has been proposed as one solution to enhance the use of guidelines in influencing standard clinical care. However, the conversion of text guidelines to the format required by a computer program is a major barrier. Clinicians who best understand the content of CPGs are typically ill equipped to convert textual guidelines into a computer accessible format. The potential of knowledge acquisition tools to assist in this process has been documented in the literature. In this paper we describe an application prototype, the Guideline Entry Wizard, created to assist in the conversion of text CPGs to a structured format within a relational database. We have tested this application through the input of information from several CPG. The application is a prototype for a more advanced tool. We have used this prototype to enter several CPGs and have demonstrated its effectiveness in inputting guideline content into a knowledge base.

INTRODUCTION

Changes in health care in the last decade have focused increased emphasis on standardization of care in order to reduce costs and improve quality. Clinical practice guidelines (CPG) can aid in standardizing care,[1] and correspondingly, there has been a significant increase in the number of published guidelines. In 1995, one study[2] reported as many as 1600 CPGs, a number too overwhelming for any individual clinician to effectively integrate into his or her practice using manual processes. The literature provides evidence for the effectiveness of computerization of CPGs for increasing compliance with recommended standards[3] and improving patient outcomes.[4,5] Although knowledge representation schemes and database models have been demonstrated for the computerization of CPGs,[6,7] computer-aided knowledge acquisition for CPGs is not as well established.[8,9] Entry of guideline content and logic is cumbersome and typically requires special training. By contrast, CPGs will likely be used more readily if their content can easily be entered into a computer system by clinicians.

In this project we have developed and implemented a software tool for knowledge acquisition to facilitate the entry of CPG into a relational database without requiring programming, specialized training, or knowledge of the underlying database. Our knowledge acquisition tool guides the collection of the content and logic required in a knowledge base for decision support. The goals of this project were to develop and evaluate a knowledge acquisition tool to assist in the conversion of text CPGs to a standardized relational format by clinicians. In this paper we describe our design approach; the challenges faced following this approach; the function of the Guideline Entry Wizard in its present phase of development; and the evaluation of the usability of the Wizard. We also discuss lessons we learned that will guide the next phase of development.

METHOD

Clinical Practice Guideline Knowledge Base

The work described in this paper is part of a larger project known as SIEGFRIED (System for Interactive Electronic Guidelines with Feedback and Resources for Instructional and Educational Development). The ultimate goal of the SIEGFRIED project is to provide interactive, guideline-based decision support on the Internet using a central "guideline server." The guideline server and associated software will permit interactive review and execution of CPGs at the point of care. An important part of the SIEGFRIED project is the use of a relational database model for the creation of a knowledge base of CPGs in a common representational format.

The guideline server database is being implemented as an open database using a relational database server (MS SQL Server, Microsoft, Inc., Redmond, WA). We have built a prototype of the database using Microsoft Access. Access was selected because it was available to the development team and the resulting database tables are portable to the SQL server. The database schema presently consists of 24 tables and 85 unique attributes. An entity relationship diagram of the database is shown in Figure 1. The design and implementation of the guideline knowledge base is described in greater detail elsewhere.[10]

1091—8280/97/$5.00

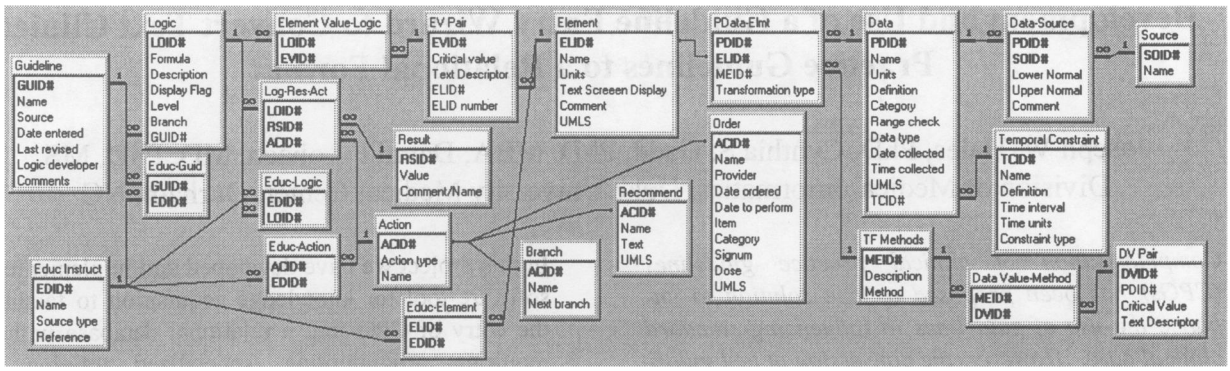

Figure 1 Entity Relationship Diagram of SIEGFRIED Knowledge Base (see Lobach[10])

Conceptual Design and Justification

One goal of the SIEGFRIED development effort is to provide clinicians with tools to encourage the use of CPGs. We hypothesize that enabling clinicians to computerize CPGs will foster increased use. Therefore, we are developing a front-end tool for entry of guidelines into the guideline server database by clinicians. The need for a data acquisition tool to facilitate the conversion of text CPGs is motivated by several factors. First, the process of transforming guidelines from text or flow charts to the relational database representation is nontrivial. Direct entry of CPG knowledge in the relational database is tedious and error-prone because of the many tables and the many relationships between tables. Important or essential information can easily be overlooked or left out. Second, a tool is needed that will assist clinical users who understand the content of the guidelines, but likely do not understand the standardized structure of the relational representation of the guidelines to input guidelines. A tool is also needed to assist the user in collecting the information required to properly complete the tables. Such a tool should not only ensure that all required fields are completed, but collect the information in a form that is consistent with the manner in which the user comprehends the guidelines and transform that information into the internal representational format.

The development of the CPG entry tool is planned in two phases. The first phase is the creation of a basic data entry program to assist knowledge engineers in the accumulation of guideline content to populate the guideline server database during prototyping of the guideline server. The tool developed during the first phase is intended to address two specific design constraints. The first is to provide rapidly a knowledge acquisition tool that will reduce errors in guideline entry. The second is to be flexible in its implementation in order to support changes to the standardized relational structure as the project evolves and as a consequence of entering real guideline information.

Using the first phase as a foundation, the second phase of development is the creation of a data entry tool that will permit a clinician that is unfamiliar with databases to enter the guideline content information. While ensuring the entry of all essential information, the tool will provide a user interface that has been abstracted from the relational representation of the database and collects the information in a method that is more consistent with the organization of the guidelines. The final tool is constrained by different factors. First, the final application must be very robust and tolerant of user errors. Second, the final application, in addition to hiding the relational structure of the underlying knowledge base, must provide substantial guidance to the user to assist in the transformation of CPGs from their textual form. The first phase has been completed and we are presently evaluating the results in preparation for the second phase.

Creation of the Guideline Entry Wizard

The design of the guideline entry program is modeled after the "wizard" concept found in many windows programs for the personal computer. A wizard is a sequence of modal dialogs that lead the user through the steps required to collect necessary information to complete a given task. The wizard masks from the user the actual location and format for storage of data collected, as well as manipulation of the data. Our wizard, called the Guideline Entry Wizard, leads the user through the steps required to create an instance for a guideline represented in our relational format.

The Guideline Wizard is built using the native forms and macro capabilities of Microsoft Access. A data entry form for each table was created using the Access form painting tools. These forms are linked sequentially by macros. Pick lists for items are derived from the list of foreign keys already in the da-

tabase through referential links to the associated tables. Additional information derived from items already entered, such as default values, is automatically loaded into subsequent forms. In cases of many-to-many relationships between tables, junction tables are required. For new, initial entries, these junction relationships are created automatically by the wizard. The execution of the guideline wizard is described below and follows the flow chart in *Figure 2*.

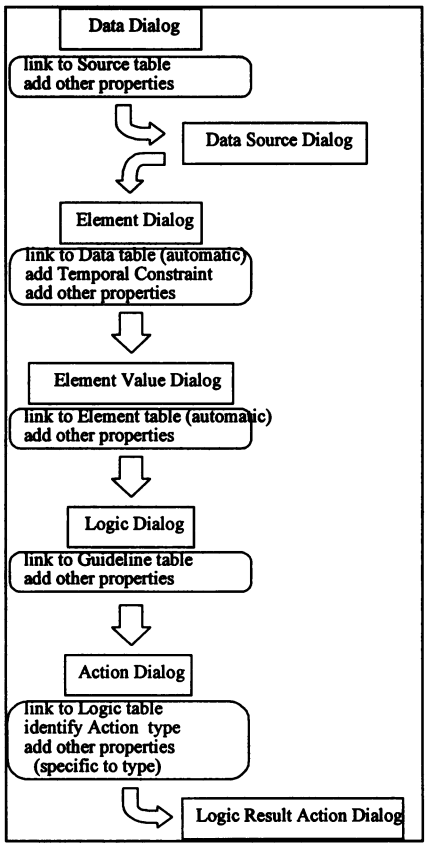

Figure 2 Flow chart showing process of entering guideline content. Rectangles represent wizard dialogs corresponding to a relational table. Rounded rectangles contain actions associated with each dialog performed by wizard. Rectangles offset to the right correspond with junction tables that create many-to-many links.

Analysis of the Guideline Entry Process

The flow of the Guideline Entry Wizard is modeled after the actual flow of content entry observed in users. Working with a clinician user (DFL) during knowledge acquisition sessions, the flow diagram in Figure 2 was developed to represent the process used to enter guideline content. Rapid prototyping was employed to provide users with a working application after the first session. Through iterative refinement working with a second knowledge engineer (CSG), the flow was further formalized and additional design requirements were added.

The overall flow of the wizard is a series of transformations of parameters (such as clinical findings) that must be collected to enable the computer to interpret the guideline logic. This process links the parameter to the guideline logic and its associated recommendations and educational resources. In *Figure 2* each rectangle represents an individual wizard dialog and corresponds to a table. Items listed in the round-edged boxes are the tasks performed as part of each dialog.

The user begins the process of entering a guideline by identifying a required parameter (referred to as a data item). Data items represent context independent observations and must be (1) transformed to a standard representation and (2) linked to the context of the guideline. The standard representation consists of conversion of the range of values of the data item to a standard scale. The data item may also be linked to a temporal transformation. Context is established by linking the parameter to a specific result value. The element-value pair is associated with a logic statement that combines the result with a procedural action if the condition of the guideline step is satisfied. Actions may be of the form of branches to additional logic, orders, or recommendations. Each of these steps is completed through a separate dialog that permits the user to select the transformation or link from a list of existing entries or to add a new entry.

RESULTS

The first phase implementation of the Guideline Entry Wizard was completed in about two weeks. The program has been used for entering portion so of two guidelines (Acute Low Back Problems in Adults from the Agency for Health Care Policy and Research and Treatment of Hypercholesterolemia from New York Life-Sanus). Prior to the completion of the prototype wizard, CPGs were entered into the database manually, directly into the tables. This manual process proved to be intractable and resulted in a number of errors in the tables. The wizard succeded in making the process of entering CPGs manageable and reduced errors in the tables because it masked the difficulty of matching coded entries in the database with their corresponding text and thus achieving the first design goal. Perhaps more importantly, the wizard followed the pattern of data entry specified by the user. Thus entry proceeded accord-

ing to a logical sequence guided by the tool leading to more complete entry of information.

Design Challenges

Several specific design challenges were faced. These design challenges can be divided into two broad categories: those that result from the approach to implementation and those that are problems associated with the domain or knowledge acquisition generally. The former are typically addressed by more or better programming, whereas, the latter represent some of the larger issues of knowledge acquisition. The latter are briefly touched on in this section and more fully addressed in the discussion.

The most substantial design challenge was the very purpose for which the Wizard was created, the translation of guidelines into the internal format of the guideline server. In practice, the text guidelines have to be carefully restructured in order to be mapped to the format of the guideline server knowledge base. The present wizard does not provide any assistance toward this preparation of the CPG. This is due in part to the decision to build a prototype quickly that could immediately be used to enter guideline content in order to keep other portions of the project on schedule. The Access macro language facilitated this rapid development; however, the macro language does not provide sufficient richness for the functions anticipated in the more advanced interface now being developed. Therefore, the present wizard does not create a foundation, as had been hoped, for the next level, only an intermediate solution.

Encoding often complicates interaction with database content. Although facilities are available in Access for translation that permits the user to deal only with the text associated with a code, the codes themselves contain no meaning. As a consequence detecting duplicately entered terms or concepts is not easily handled and requires additional programming. Another shortcoming of the code handling ability of Access is in the construction of the actual guideline logic formulae. As noted in Figure 3, the logic formulae contain codes for element-value pairs. Ideally, a logic statement builder could be programmed that would graphically assist in the creation of the logic formula and provide the necessary translation of codes, as well.

A second important design challenge was the translation of the process of guideline entry. A wizard leads one through a sequence of steps necessary to complete a task, in this case, entry of a guideline. In reality, the process of creating a guideline does not proceed in a linear, step-wise fashion, but entails

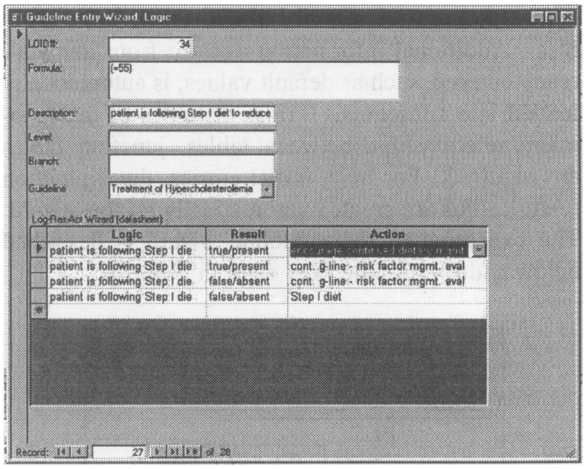

Figure 3. A screen shot of the Logic dialog. The upper portion contains the slots for a single logic statement. Note the code of an element-value pair in the formula field. The lower portion of the screen contains the relationship of this logic statement to various outcomes and associated actions.

many iterative sub-steps in preparation to the formulation of the actual logic statement. Further, the sequence imposed by the wizard is not the only reasonable sequence for entering guideline information. For example, one might start at the guideline level, moving to the logic level and at that point identify the data elements that are required. This approach, along with other reasonable sequences or partial sequences, should be supported.

DISCUSSION

In this project we have created a Guideline entry Wizard that provides a front end to a relational database containing a knowledge base for CPGs. The knowledge acquisition tool we have created has successfully addressed the design constraints identified, that of being developed rapidly and reducing errors in the entry of guideline content.

As noted previously, a number of design challenges were exposed during the first phase of implementation. By design, many of these will be addressed in the second phase of development. However, several of the limitations have emerged as a result of the creation of the application. First, the flow and input format of the Guideline Entry Wizard is presently tightly coupled to the underlying relational format in which the data are stored. This tight coupling is by design, as noted above; however, this aspect has required that the user entering guideline information (in this case, one of the designers of the database schema; CSG) be fully aware of the relational repre-

sentation of the guidelines. A related consequence is the impact on the application of changes to the relational schema which proved more troublesome than expected in updating the Guideline Entry Wizard. Although the tight coupling of the data entry screens to the actual tables is considered a disadvantage, the modularity permitted some reuse of macros.

Another important limitation is the scalability of the data entry mechanisms selected. We anticipate as the database grows that the utility of the pick list mode of entry will decline because the size of the pick lists will become intractable. The native tools of Access provide only limited alphabetical sorting by first word in a phrase for the pick lists. Much more sophisticated capabilities will need to be programmed into the guideline entry application in order to take advantage of reuse of the encoded data.

Two limitations have emerged as a result of entering guidelines and in connection with proceeding with other development of the Guideline Server. One limitation is the absence of a cumulative view that would support the display of the information in the database for one guideline in a manner consistent with the way the interactive display engine will interact with the data. Presently, the actual processing of guidelines entered must be tracked manually through paper copies of the database. This is as tedious as manual entry and prone to substantial errors. An automated tool, independent of the interactive guideline application, to similarly process the guideline information would prove very useful. Although not part of the original design for phase two, this capability is considered an essential addition to the development of the second phase application.

A second limitation that has emerged as part of the process has been the absence of a graphical presentation of the guideline information, either once entered, or as a method of entry. Many of the guidelines originate in a graphical form, typically a flow chart. Some consideration has been given to the utility of permitting the user to enter the guideline information into the system in graphical form or to at least review entered guidelines in a graphical form. One limitation to the graphical entry of guideline information is the need for the user to learn the graphical data entry tool. The difficulty of this has not been assessed at this time and must be balanced against the increased complexity of the guideline entry application.

Future work will focus first on addressing as part of the second phase those design issues discussed above. The Guideline Entry Wizard also provides a mechanism for better understanding the cognitive process of the clinician during transformation of text guidelines to a computerized form.

Acknowledgments

This work supported in part by grant R01 HS09436-01 from the Agency for Health Care Policy and Research.

References

[1] Audet AM, Greenfield S. Field M. Medical practice guidelines: current activities and future directions. Ann Intern Med. 1990;;113:709-714.

[2] American Medical Association. Directory of Practice parameters. Chicago, Ill.: American Medical Association; 1995.

[3] Lobach DF, Hammond WE. Computerized decision support based on clinical practice guideline improves compliance with care standards. Amer J Med. 1997;102:89-98.

[4] Grimshaw JM, Russell IT. Effect of clinical practice guidelines on medical practice: a systematic review of rigorous evaluations. Lancet. 1993;342:1317-1322.

[5] Committee on Clinical Practice Guidelines, Institute of Medicine. Guidelines for Clinical Practice: From Development to Use. Field MJ, Lohr KN (eds). Washington DC: National Academy Press, 1990.

[6] ASTM Subcommittee E31.15 on Medical Knowledge Representation. Standard specificatoin for defining and sharing modular health knowledge bases (Arden Syntax for ledical logic systems). Philadelphia, PA: ASTM; 1992.

[7] Kuperman GJ, Teich JM, Bates DW, McLatchey J, Hoff TG. Representing hospital events as complex conditionals. In: Gardner RM, ed. 19th SCAMC. Philadelphia: Hanley & Belfus, Inc. 1995:137-141.

[8] Musen MA, Tu SW, Das AK, Shahar Y. EON: A component-based approach to automation of protocol-directed therapy. JAMIA. 1996;3:367-388.

[9] Stempczynska J, Kacki E. Problems of knowledge acquisition automation in medical expert systems. In Greenes RA, et. al., ed., Proceedings of MEDINFO '95. Vancouver, British Columbia: IMIA. 1995;857-860.

[10] Lobach DF, Gadd CS, Hales JW. Structuring Diverse Clinical Practice Guidelines in a Relational Database Model for Decision Support on the Internet. 1997 AMIA Annual Fall Symposium.